

# **LOOK-UP TABLE BASED LOGIC MACRO-CELLS**

**Inventor: R. U. Madurawe**

This application is related to Application Serial No. 10/267,484 entitled "Methods  
5 for Fabricating Three-Dimensional Integrated Circuits", Application Serial No.  
10/267,483 entitled "Three Dimensional Integrated Circuits", and Application Serial No.  
10/267,511 entitled "Field Programmable Gate Array With Convertibility to Application  
Specific Integrated Circuit", all of which were filed on October 8, 2002 and list as  
inventor Mr. R. U. Madurawe, the contents of which are incorporated herein by  
10 reference.

This application is also related to Application Serial Number 10/413,809 entitled  
"Semiconductor Switching Devices", Application Serial No. 10/413,808 entitled  
"Insulated-Gate Field-Effect Thin Film Transistors", and Application Serial No.  
10/413,810 entitled "Semiconductor Latches and SRAM Devices", all of which were  
15 filed April 14, 2003 and list as inventor Mr. R. U. Madurawe, the contents of which are  
incorporated herein by reference.

## **BACKGROUND**

The present invention relates to look up table based macrocells for programmable  
20 logic applications.

Traditionally, application specific integrated circuit (ASIC) devices have been  
used in the integrated circuit (IC) industry to reduce cost, enhance performance or meet  
space constraints. The generic class of ASIC devices falls under a variety of sub classes

such as Custom ASIC, Standard cell ASIC, Gate Array and Field Programmable Gate Array (FPGA) where the degree of user allowed customization varies. In this disclosure the word ASIC is used only in reference to Custom and Standard Cell ASICs where the designer has to incur the cost of a full fabrication mask set. The term FPGA denotes an off the shelf programmable device with no fabrication mask costs, and Gate Array denotes a device with partial mask costs to the designer. The devices FPGA include Programmable Logic Devices (PLD) and Complex Programmable Logic Devices (CPLD), while the devices Gate Array include Laser Programmable Gate Arrays (LPGA), Mask Programmable Gate Arrays (MPGA) and a new class of devices known as Structured ASIC or Structured Arrays.

The design and fabrication of ASICs can be time consuming and expensive. The customization involves a lengthy design cycle during the product definition phase and high Non Recurring Engineering (NRE) costs during manufacturing phase. In the event of finding a logic error in the custom or semi-custom ASIC during final test phase, the design and fabrication cycle has to be repeated. Such lengthy correction cycles further aggravate the time to market and engineering cost. As a result, ASICs serve only specific applications and are custom built for high volume and low cost. The high cost of masks and unpredictable device life time shipment volumes have caused ASIC design starts to fall precipitously in the IC industry. ASICs offer no device for immediate design verification, no interactive design adjustment capability, and require a full mask set for fabrication.

Gate Array customizes pre-defined modular blocks at a reduced NRE cost by designing the module connections with a software tool similar to that in ASIC. The Gate

Array has an array of non programmable (or moderately programmable) functional modules fabricated on a semiconductor substrate. To interconnect these modules to a user specification, multiple layers of wires are used during design synthesis. The level of customization may be limited to a single metal layer, or single via layer, or multiple metal layers, or multiple metals and via layers. The goal is to reduce the customization cost to the user, and provide the customized product faster. As a result, the customizable layers are designed to be the top most metal and via layers of a semiconductor fabrication process. This is an inconvenient location to customize wires. The customized transistors are located at the substrate level of the Silicon. All possible connections have to come up to the top level metal. The complexity of bringing up connections is a severe constraint for these devices. Structured ASICs fall into larger module Gate Arrays. These devices have varying degrees of complexity in the structured cell and varying degrees of complexity in the custom interconnection. The absence of Silicon for design verification and design optimization results in multiple spins and lengthy design iterations to the end user. The Gate Array evaluation phase is no different to that of an ASIC. The advantage over ASIC is in a lower upfront NRE cost for the fewer customization layers, tools and labor, and the shorter time to receive the finished product. Gate Arrays offer no device for immediate design verification, no interactive design adjustment capability, and require a partial mask set for fabrication. Compared to ASICs, Gate Arrays offer a lower initial cost and a faster turn-around to debug the design. The end IC is more expensive compared to an ASIC.

In recent years there has been a move away from custom, semi-custom and Gate Array ICs toward field programmable components whose function is determined not

when the integrated circuit is fabricated, but by an end user “in the field” prior to use. Off the shelf FPGA products greatly simplify the design cycle and are fully customized by the user. These products offer user-friendly software to fit custom logic into the device through programmability, and the capability to tweak and optimize designs to improve

5 Silicon performance. Provision of this programmability is expensive in terms of Silicon real estate, but reduces design cycle time, time to solution (TTS) and upfront NRE cost to the designer. FPGAs offer the advantages of low NRE costs, fast turnaround (designs can be placed and routed on an FPGA in typically a few minutes), and low risk since designs can be easily amended late in the product design cycle. It is only for high volume

10 production runs that there is a cost benefit in using the other two approaches. Compared to FPGA, an ASIC and Gate Array both have hard-wired logic connections, identified during the chip design phase. ASIC has no multiple logic choices and both ASIC and most Gate Arrays have no configuration memory to customize logic. This is a large chip area and a product cost saving for these approaches to design. Smaller die sizes also lead

15 to better performance. A full custom ASIC has customized logic functions which take less gate counts compared to Gate Arrays and FPGA configurations of the same functions. Thus, an ASIC is significantly smaller, faster, cheaper and more reliable than an equivalent gate-count FPGA. A Gate Array is also smaller, faster and cheaper compared to an equivalent FPGA. The trade-off is between time-to-market (FPGA

20 advantage) versus low cost and better reliability (ASIC advantage). A Gate Array falls in the middle with an improvement in the ASIC NRE cost at a moderate penalty to product cost and performance. The cost of Silicon real estate for programmability provided by the

FPGA compared to ASIC and Gate Array contribute to a significant portion of the extra cost the user has to bear for customer re-configurability in logic functions.

In an FPGA, a complex logic design is broken down to smaller logic blocks and programmed into logic blocks provided in the FPGA. Logic blocks contain multiple smaller logic elements. Logic elements facilitates sequential and combinational logic design implementations. Combinational logic has no memory and outputs reflect a function solely of present input states. Sequential logic is implemented by inserting memory in the form of a flip-flop into the logic path to store past history. Current FPGA architectures include transistor pairs, NAND or OR gates, multiplexers, look-up-tables (LUT) and AND-OR structures in a basic logic element. In a PLD the basic logic element is labeled a macro-cell. Hereafter the terminology logic element will include both logic elements and macro-cells. Granularity of an FPGA refers to logic content in the basic logic block. Partitioned smaller blocks of a complex logic design are customized to fit into FPGA grain. In fine-grain architectures, one or a few small basic logic elements are grouped to form a basic logic block, then enclosed in a routing matrix and replicated. A fine grain logic element may contain a 2-input MUX or a 2-input LUT and a register. These offer easy logic fitting at the expense of complex routing. In course grain architectures, many larger logic elements are combined into a basic logic block with local routing. A course grain logic element may include a 4-input LUT with a register, and a logic block may include as many as 4 to 8 logic elements. The larger logic block is then replicated with a global routing matrix. Larger logic blocks make the logic fitting difficult and the routing easier. A challenge for FPGA architectures is to provide easy logic fitting (like fine grain) and maintain easy routing (like course grain). Course grain architectures

are faster in logic operations and there is an increasing need in the IC industry to utilize larger logic blocks with multiple bigger LUT structures.

For sequential logic designs, the logic element may also include flip-flops. A MUX based exemplary logic element described in Ref-1 (Seals & Whapshott) is shown in Fig-1A. The logic element has a built in D-flip-flop 105 for sequential logic implementation. In addition, elements 101, 102 and 103 are 2:1 MUX's controlled by one input signal for each MUX. Input S1 feeds into 101 and 102, while inputs S1 and S2 feeds into OR gate 104, and the output from OR gate feeds into 103. Element 105 is the D-Flip-Flop receiving Preset, Clear and Clock signals. One may very easily represent the programmable MUX structure in Fig-1A as a 2-input LUT; where A, B, C & D are LUT values, and S1, (S2+S3) are LUT inputs. Ignoring the global Preset & Clear signals, eight inputs feed into the logic block, and one output leaves the logic block. All 2-input, all 3-input and some 4-input variable functions are realized in the logic block and latched to the D-Flip-Flop. Inputs and outputs for the Logic Element or Logic Block are selected from the programmable Routing Matrix. An exemplary routing matrix containing logic elements as described in Ref-1 is shown in Fig-1B. Each logic element 112 is as shown in Fig-1A. The 8 inputs and 1 output from logic element 112 in Fig-1B are routed to 22 horizontal and 12 vertical interconnect wires that have programmable via connections 110. These connections 110 may be anti-fuses or pass-gate transistors controlled by SRAM memory elements. The user selects how the wires are connected during the design phase, and programs the connections in the field. FPGA architectures for various commercially available FPGA devices are discussed in Ref-1 (Seals & Whapshott) and Ref-2 (Sharma).

Logic implementation in logic elements is achieved by converting a logic equation or a truth table to a gate realization. The gate level description comprising elements and nets is also called a netlist. The resulting logic gates are ported to LUT or MUX structure in the logic element. An exemplary truth table and a plurality of transistor gate realizations are shown in Fig-2. In Fig-2A, a truth table of 4 input variables, A, B, C & D is shown. By grouping the logic ones in the table, the output function can be expressed as AND & OR functions of inputs as shown by the logic equation in Fig-2A. An exemplary MUX implementation of the logic function is shown in Fig-2B. The MUX has 3-control variables A, B and C, and the fourth variable D together with D' (not D), logic one and logic zero are used as inputs to the MUX. The inputs can be hard-wired or provided as programmable options. The MUX comprises a plurality of pass-gates 201. For a 3-variable hard-wired MUX, only 14 pass-gates such as 201 are needed. This is a very efficient implementation of hard-wired logic. Any 4-variable truth table can be realized by the 3-control variable MUX as shown in Fig-2B by wiring the input values accordingly. The inputs to a programmable MUX logic element can be provided as shown in Fig-2C. There is considerable overhead to make the MUX inputs user programmable. In Fig-2C, two programmable memory bits such as 202 per input are configured to couple the desired input value to I<sub>1</sub>. Combining the two figures in Fig-2B & 2C, one can see that a 4-input programmable MUX utilizes 62 pass-gates such as 201 and 16 memory bits such as 202. For 6T CMOS SRAM memory, each memory bit occupies 4 NMOS gates and 2 PMOS gates. Hence a programmable 4-input MUX implementation takes up 158 transistors. In anti-fuse technology, each input wire connection can be built into a programmable anti-fuse between two metal lines. That requires only decoding

transistors at the end of wire segments to program the anti-fuse elements, thus saving Silicon area. Hence a programmable MUX as shown in Fig-2B is not popular for SRAM based FPGAs, whereas it is a logical choice for anti-fuse based FPGAs.

AND/OR realization of the logic function in Fig-2A is shown in Fig-2D. There are five 3-input AND gates and one 5-input OR gate to generate the required F output. In full CMOS implementation, each 3-input AND is 6 transistors, while 5-input OR is 10 transistors. Hence the AND/OR gate realization in Fig-2D takes up 40 transistors. The Silicon area is also impacted by the latch-up related N-Well rules that mandate certain spacing restrictions between NMOS and PMOS transistors. For this example, the hard-wire MUX implementation took less gates compared to the hard-wire AND/OR gate implementation, while the programmable MUX took a considerable overhead.

Commercially available FPGAs use 3-input and 4-input look up tables (LUT). The more popular 4-input LUT implementation of the truth table in Fig-2A is shown in Fig-2E. Any 4-input function can be implemented in Fig-2E by setting the LUT values. In this disclosure, we will name this a 4LUT, where the word input is dropped for convenience and the number of inputs is pre-fixed to the word LUT. The 4LUT has 16 LUT values, which can be hard-wired or programmable. LUT and MUX construction of logic elements are very similar and both are commercially used in FPGA & Gate Array products as shown in Ref-1 & Ref-2. There are 30 pass-gates (such as 201) in Fig-2E for the hard-wire 4LUT. This 30 gate 4LUT is larger than a 14 gate hard-wire MUX, but smaller than the 40 gate hard-wire AND/OR logic implementation. The 16 LUT values in the 4LUT determine the LUT function. Using 16 programmable registers such as 202 for these inputs allows the 4LUT to be user programmable. The 16 memory elements, in both



programmable MUX and LUT options, utilize 96 extra transistors when implemented in 6T CMOS SRAM. Hence the programmable 4LUT with 126 transistors is more economical compared to the programmable MUX option with 158 transistors. Thus LUT logic is extensively used in SRAM based FPGAs while MUX logic is used in anti-fuse based FPGAs and Gate Arrays.

FPGA and Gate Array architectures are discussed in Carter US 4706216, Freemann US 4870302, ElGamal et al. US 4873459, Freemann et al. US 5488316 & US 5343406, Trimberger et al. US 5844422, Cliff et al. US 6134173, Wittig et al. US 6208163, Or-Bach US 2001/003428, Mendel US 6275065, Lee et al. US 2001/0048320, Or-Bach US 6331789, Young et al. US 6448808, Sueyoshi et al. US 2003/0001615, Agrawal et al. US 2002/0186044, Sugibayashi et al. US 6515511 and Pugh et al. US 2003/0085733. These patents disclose programmable MUX and programmable LUT structures to build logic elements that are user configurable. In all cases a routing block is used to provide inputs and outputs for these logic elements, while the logic element is programmed to perform a specific logic function. The routing-block is a hard-wire connection for Gate Array and Structured ASIC devices. Within a logic element, each LUT is hard-wired to a specific size, said size determined by the number of LUT inputs. This LUT is the smallest building block in the logic element and cannot be sub-divided. As an example, a smaller 2-input logic function would occupy a 4LUT, if that is the smallest element available. That leads to Silicon utilization inefficiency. Within a logic block, multiple logic elements are grouped together in a pre-defined manner. The size of the logic block determines the granularity. As manufacturing geometries shrink, the FPGA granularity gets larger, the LUT size increases and the number of LUTs per logic

block has to increase. Having a large fixed LUT in the logic element further aggravates the Silicon utilization efficiency and is not flexible for next generation FPGA designs.

As the LUT structure gets large, the logic porting becomes more difficult and Silicon utilization gets more inefficient. To illustrate LUT utilization efficiency, in Fig-3 we provide the pass-gate construction required to build 1LUT, 2LUT, 3LUT, 4LUT and 5LUT logic elements. Fig-3A shows a 1LUT comprising of two pass-gates 301 & 302, two LUT values contained in two programmable registers 303 & 304 and one input variable "A" in true and compliment. A 1LUT is simply a 2:1 MUX selecting one of two register values. Any 1-input function such as 2:1 MUX, Logic-1, Logic-0, TRUE and INVERT can be realized by this 1LUT by programming the two LUT values. Signal A allows the LUT values in either 303 or 304 to reach output F. There is a time delay for this to occur. That is a characteristic 1LUT delay time, which is optimized by sizing the transistors 301 and 302 as needed. Faster time requires wider transistors. The symbol for 1LUT is shown in Fig-3B, and this symbol is used to illustrate higher LUT constructions in Fig-3C thru Fig-3F.

A 2LUT is shown in Fig-3C that can realize any 2-input function such as AND, NAND, OR, NOR, XOR among others. As shown in Fig-3C, the 2LUT can be constructed by hard-wiring three 1LUTs 311, 312 & 313 as shown. This is termed a LUT cone or a LUT tree and comprises two stages. First stage has 1LUT 311 and 312 sharing a common input, while second stage has 1LUT 313. Only the 1LUTs in the first stage 311 and 312 have LUT values. LUT outputs from first stage are fed as LUT values to second stage. These are hard-wire connections. In Fig-3C, 1LUT outputs from 311 and 312 are fed as LUT values to 1LUT 313. A 2LUT delay comprises the time taken for a

LUT value in the first stage to reach F. There are now two pass-gates in series, and this delay is larger than for a 1LUT. Thus the pass-gates need to be wider to reduce the LUT delay. That increase in area and slow down in performance hurt LUT logic trees. Similarly, 3LUT, 4LUT and 5LUT constructions with 1LUTs are shown in Fig-3D, Fig-3E and Fig-3F respectively. Those pass-gates have to be even wider to improve LUT delays. The 5LUT in Fig-3F has 16 1LUTs in the first stage, 8 1LUTs in the second stage, 4 1LUTs in the third stage, 2 1LUTs in the fourth stage and one 1LUT in the final fifth stage. A total of 31 1LUTs are used in Fig-3F for the 5LUT construction. A K-LUT cone or a K-LUT tree has K-input variables, K-stages and  $2^K$  LUT values to realize a K-input function. Each stage has one common input variable.  $2^{(K-1)}$  outputs from first stage feed as LUT values into second stage. Consecutive LUT value reduction continues until the last stage, when only 2 LUT values feed the last stage, and one LUT output is obtained. The equivalent 1LUTs required to build a K-LUT is tabulated in Fig-3G, and is shown to grow as  $(2^K-1)$ . Logic porting to K-LUT is discussed by Ahmed et al. (Ref-3) for multiple K values. They have looked at porting 20 benchmark logic designs into varying LUT sizes: 1LUT, 2LUT, 3LUT, 4LUT, 5LUT, 6LUT and 7LUT. The geometric average number of K-LUTs required for porting 20 designs, as shown in Fig-10 in Ref-2, is tabulated in the first 2 columns of Fig-4. As can be seen, as the size of the K-LUT increases, the total number of K-LUTs required to fit an average design decreases. In addition, Fig-4 also lists the equivalent 1LUT per K-LUT (from Fig-3G) in column 3, and calculates the equivalent 1LUTs required for the design in column 4. Column 4 values are obtained by multiplying values in column 2 by values in column 3. In Fig-4, each row represents how many K-LUTs are required for an average design, and an equivalent

1LUT calculation as a measure of Silicon utilization. 2LUT implementation in row-1 needs only 12900 1LUTs, while the 7LUT implementation in row-6 needs 177800 1LUTs for the same design. The latter 7LUT has only 7.3% Silicon utilization efficiency compared to the former 2LUT. From row-3, commercially available FPGAs with 4LUTs are seen only 36.1% efficient compared to 2LUTs at fitting logic. As the LUT size gets larger, clearly a more efficient LUT circuit is needed to improve Silicon utilization in LUT based logic elements.

LUT based logic elements are used in conjunction with programmable point to point connections. Four exemplary methods of programmable point to point connections, synonymous with programmable switches, between node A and node B are shown in Fig-5. A configuration circuit to program the connection is not shown in Fig-5. All the patents listed under FPGA architectures use one or more of these basic programmable connections. In Fig-5A, a conductive fuse link 510 connects A to B. It is normally connected, and passage of a high current or exposure to a laser beam will blow the conductor open. In Fig-5B, a capacitive anti-fuse element 520 disconnects A from B. It is normally open, and passage of a high current will pop the insulator shorting the two terminals. Fuse and anti-fuse are both one time programmable due to the non-reversible nature of the change. In Fig-5C, a pass-gate device 530 connects A to B. The gate signal  $S_0$  determines the nature of the connection, on or off. This is a non destructive change. The gate signal is generated by manipulating logic signals, or by configuration circuits that include memory. The choice of memory varies from user to user. In Fig-5D, a floating-pass-gate device 540 connects A to B. Control gate signal  $S_0$  couples a portion of that to floating gate. Electrons trapped in the floating gate determines an on or off state

for the connection. Hot-electrons and Fowler-Nordheim tunneling are two mechanisms for injecting charge to floating-gates. When high quality insulators encapsulate the floating gate, trapped charge stays for over 10 years. These provide non-volatile memory. EPROM, EEPROM and Flash memory employ floating-gates and are non-volatile. Anti-fuse and SRAM based architectures are widely used in commercial FPGA's, while EPROM, EEPROM, anti-fuse and fuse links are widely used in commercial PLD's. Volatile SRAM memory needs no high programming voltages, is freely available in every logic process, is compatible with standard CMOS SRAM memory, lends to process and voltage scaling and has become the de-facto choice for modern day very large FPGA device construction.

All commercially available high density FPGA's use SRAM memory elements. A volatile six transistor SRAM based configuration circuit is shown in Fig-6A. The SRAM memory element can be any one of 6-transistor, 5-transistor, full CMOS, R-load or TFT PMOS load based cells to name a few. Two inverters 603 and 604 connected back to back forms the memory element. This memory element is a latch providing complementary outputs  $S_0$  and  $S_0'$ . The latch can be constructed as full CMOS, R-load, PMOS load or any other. Power and ground terminals for the inverters are not shown in Fig-6A. Access NMOS transistors 601 and 602, and access wires GA, GB, BL and BS provide the means to configure the memory element. Applying zero and one on BL and BS respectively, and raising GA and GB high enables writing zero into device 601 and one into device 602. The output  $S_0$  delivers a logic one. Applying one and zero on BL and BS respectively, and raising GA and GB high enables writing one into device 601 and zero into device 602. The output  $S_0$  delivers a logic zero. The SRAM construction

may allow applying only a zero signal at BL or BS to write data into the latch. The SRAM cell may have only one access transistor 601 or 602. The SRAM latch will hold the data state as long as power is on. When the power is turned off, the SRAM bit needs to be restored to its previous state from an outside permanent memory. In the literature for programmable logic, this second non-volatile memory is also called configuration memory. Upon power up, an external or an internal CPU loads the external configuration memory to internal configuration memory locations. All of FPGA functionality is controlled by the internal configuration memory. The SRAM configuration circuit in Fig-6A controlling logic pass-gate is illustrated in Fig-6B. Element 650 represents the configuration circuit. The  $S_0$  output directly driven by the memory element shown in Fig-6A drives the pass-gate 610 gate electrode. In addition to  $S_0$  output and the memory cell, power, ground, data-in and write-enable signals in 650 constitutes the SRAM configuration circuit. Write enable circuitry includes GA, GB, BL, BS signals shown in Fig-6A.

As discussed earlier, providing programmability is a very severe transistor and cost penalty compared to hard-wired Gate Array or ASIC implementation of identical logic. A significant factor in the penalty comes from the 6-transistors required for the configuration circuits. The natural conclusion is to minimize the number of configurable bits used in the programmable logic element. This mandates constructing a hard-wired larger 6LUT or a bigger LUT for next generation FPGAs. We have shown that Silicon utilization is severely impacted with this move towards larger LUT structures in logic elements. What is desirable is to have an economical and flexible LUT macro-cell, or a macro-LUT circuit. This LUT macro-cell should efficiently implement logic functions.

Both large logic functions that port to one big LUT and small logic functions that port to multiple smaller LUTs should fit easily into a LUT macro-cell. Furthermore, LUT logic packing should maximize Silicon utilization to keep programmable logic cost reasonable with other hard-wired IC manufacturing choices. The user should be able to take a synthesized netlist from an ASIC flow, typically comprising smaller logic blocks, convert this netlist to fit in the FPGA granularity, place and route logic economically and efficiently. This would make use of existing third party ASIC tools at the front-end logic design and streamline tool flow for FPGA place & routing.

For an emulation device, the cost of programmability is not the primary concern if such a device provides a migration path to a lower cost. Today an FPGA migration to a Gate Array requires a new design to ensure timing closure. A desirable migration path is to keep the timing of the original FPGA design intact. That would avoid valuable re-engineering time, opportunity costs and time to solution (TTS). Such a conversion should occur in the same base die to avoid Silicon and system re-qualification costs and implementation delays. Such a conversion should also realize an end product that is competitive with an equivalent standard cell ASIC or a Gate Array product in cost and performance. Such an FPGA device will also target applications that are cost sensitive, have short life cycles and demand high volumes.

## SUMMARY

In one aspect, a programmable look up table (LUT) circuit for an integrated circuit comprises: one or more secondary inputs; and one or more configurable logic states; and two or more LUT values; and a programmable means to select a LUT value from a secondary input or a configurable logic state.

Implementations of the above aspect may include one or more of the following. A semiconductor integrated circuit comprises an array of programmable modules. Each module may use one or more LUT or MUX based logic elements. A programmable interconnect structure may be used to interconnect these programmable modules in an FPGA device. A logic design may be specified by the user in VHDL or Verilog design input language and synthesized to a gate-level netlist description. This synthesized netlist is ported into logic blocks and connected by the routing block in the FPGA. Each large LUT in a module may be comprised of a smaller 1-input LUT (1LUT) cone, known also as a 1LUT tree. A Larger LUT may be comprised of smaller 2LUT, or 3LUT trees. A smaller LUT provides added flexibility in fitting logic. A smaller LUT provides at least one LUT value to be selected from either a programmable register or from an input. The input may be an output of a previously generated logic function, or an external input. The registers may be user configurable to logic zero and logic one states. The larger LUT and smaller LUT may comprise a programmable switch to connect two points. Most common switch is a pass-gate device. A pass-gate is an NMOS transistor, or a PMOS transistor or a CMOS transistor pair that can electrically connect two points. Other methods of connecting two points include fuse links and anti-fuse capacitors, among others. Programming these devices include forming one of either a conducting path or a non-



conducting path in the connecting device. These pass-gates may be fabricated in a first module layer, said module comprising a Silicon substrate layer.

The LUT circuits may include digital circuits consisting of CMOS transistors forming AND, NAND, INVERT, OR, NOR and pass-gate type logic circuits.

5 Configuration circuits are used to change LUT values, functionality and connectivity.

Configuration circuits have memory elements and access circuitry to change stored memory data. Memory elements can be RAM or ROM. Each memory element can be a transistor or a diode or a group of electronic devices. The memory elements can be made of CMOS devices, capacitors, diodes, resistors, wires and other electronic components.

10 The memory elements can be made of thin film devices such as thin film transistors (TFT), thin-film capacitors and thin-film diodes. The memory element can be selected from the group consisting of volatile and non volatile memory elements. The memory element can also be selected from the group comprising fuses, antifuses, SRAM cells, DRAM cells, optical cells, metal optional links, EPROMs, EEPROMs, flash, magnetic,  
15 electro-chemical and ferro-electric elements. One or more redundant memory elements can be provided for controlling the same circuit block. The memory element can generate an output signal to control pass-gate logic. Memory element may generate a signal that is used to derive a control signal to control pass-gate logic. The control signal is coupled to MUX or Look-Up-Table (LUT) logic element.

20 LUT circuits are fabricated using a basic logic process used to build CMOS transistors. These transistors are formed on a P-type, N-type, epi or SOI substrate wafer. Configuration circuits, including configuration memory, constructed on same Silicon substrate take up a large Silicon foot print. That adds to the cost of programmable LUT

circuits compared to similar functionality custom wire circuits. A 3-dimensional integration of configuration circuits described in incorporated references provides a significant cost reduction in programmability. The configuration circuits may be constructed after a first contact layer is formed or above one or more metal layers. The programmable LUT may be constructed as logic circuits and configuration circuits. The configuration circuits may be formed vertically above the logic circuits by inserting a thin-film transistor (TFT) module. The TFT module may include one or more metal layers for local interconnect between TFT transistors. The TFT module may include salicided poly-Silicon local interconnect lines and thin film memory elements. The thin-film module may comprise thin-film RAM elements. The thin-film memory outputs may be directly coupled to gate electrodes of LUT pass-gates to provide programmability. Contact or via thru-holes may be used to connect TFT module to underneath layers. The thru-holes may be filled with Titanium-Tungsten, Tungsten, Tungsten Silicide, or some other refractory metal. The thru-holes may contain Nickel to assist Metal Induced Laser Crystallization (MILC) in subsequent processing. Memory elements may include TFT transistors, capacitors and diodes. Metal layers above the TFT layers may be used for all other routing. This simple vertically integrated pass-gate switch and configuration circuit reduces programmable LUT cost.

In a second aspect, a programmable look up table circuit for an integrated circuit comprises:  $M$  primary inputs, wherein  $M$  is an integer value greater than or equal to one, and each said  $M$  inputs received in true and compliment logic levels; and  $2^M$  secondary inputs; and  $2^M$  configurable logic states, each said state comprising a logic zero and a logic one; and  $2^M$  LUT values; and a programmable means to select each of said LUT

values from a secondary input or a configurable logic state.

Implementations of the above aspect may include one or more of the following. A larger N-LUT is constructed with all equal size smaller K-LUTs. A larger N-LUT is constructed with unequal sized smaller K-LUTs. Each smaller K-LUT is constructed as a 1LUT, 2LUT, 3LUT up to (N-1)-LUT. The N-LUT is constructed as a K-LUT tree. Each stage in the N-LUT tree comprises a plurality of K-LUTs. Each K-LUT has one output. Larger N-LUT has one or more outputs comprising a plurality of smaller K-LUT outputs. Each K-LUT is also constructed as a 1LUTs tree. All primary K-LUTs (the first set of K-LUTs) in the N-LUT tree may have only configurable logic states for LUT values. All primary K-LUTs may have a LUT value selected from an input and a configurable logic state. Said input may comprise an external input, a feed-back signal, a memory output or a control signal. Secondary K-LUT in the N-LUT tree provides a programmable connection between previous K-LUT outputs and configurable logic states. This hierarchical K-LUT arrangement is termed herein a LUT macrocell circuit. A LUT macrocell provides programmability to implement logic as one large N-LUT or as multiple smaller K-LUTs. Such division in logic implementation allows more logic to fit in a single LUT macrocell. It provides course-grain architecture with fine-grain logic fitting capability. More logic fitting improves Silicon utilization. In one embodiment, the smaller K-LUTs are implemented as 1LUTs. In a second embodiment the smaller K-LUTs are implemented as 2LUTs. In yet another embodiment the smaller K-LUTs are implemented as 3LUTs. A 1LUT in the first stage of a secondary K-LUT is used to combine two outputs from prior K-LUTs.

In a third aspect, a programmable macro look up table (macro-LUT) circuit for an

integrated circuit, comprises: a plurality of LUT circuits, each of said LUT circuits comprising a LUT output, at least one LUT input, and at least two LUT values; and a programmable means of selecting LUT inputs to at least one of said LUT circuits from one or more other LUT circuit outputs and external inputs, and selecting LUT values to at least one of said LUT circuits from one or more other LUT circuit outputs and configurable logic states, said programmable means further comprised of two selectable manufacturing configurations, wherein: in a first selectable configuration, a random access memory circuit (RAM) is formed, said memory circuit further comprising configurable thin-film memory elements; in a second selectable configuration, a hard-wire read only memory circuit (ROM) is formed in lieu of said RAM, said ROM duplicating one RAM pattern in the first selectable option.

Implementations of the above aspect may include one or more of the following. A programmable macro-LUT is used for a user to customize logic in an FPGA. This programmability is provided to the user in an off the shelf FPGA product. There is no waiting and time lost to port synthesized logic design into a macro-LUT circuit. This reduces time to solution (TTS) by 6 months to over a year. The macro-LUT can be subdivided into smaller LUT circuits. Each smaller LUT is comprised of 1LUTs. A portion of macro-LUT inputs and LUT values are selected by a programmable method. This allows prior LUT output logic manipulation. Macro-LUT inputs are selected from external inputs or other LUT outputs. LUT values are selected from external inputs, other LUT outputs or configurable logic states. Macro-LUT is very flexible in fitting one large logic block or many smaller logic blocks. Macro-LUT improves Silicon utilization. Macro-LUT improves run-times of a software tool that ports logic designs into FPGA.

Macro-LUT improves routability. The Macro-LUT is constructed with RAM and ROM options.

Implementations of the above aspect may include one or more of the following. A programmable method includes customizing programmable LUT choices. This may be done by the user, wherein the macro-LUT comprises configuration circuits, said circuits including memory elements. Configuration circuits may be constructed in a second module, substantially above a first module comprising LUT pass-gate transistors. Configuration memory is built as Random Access Memory (RAM). User may customize the RAM module to program the LUT connections. The RAM circuitry may be confined to a thin-film transistor (TFT) layer in the second module. This TFT module may be inserted to a logic process. Manufacturing cost of TFT layers add extra cost to the finished product. This cost makes a programmable LUT less attractive to a user who has completed the programming selection. Once the programming is finalized by the user, the LUT connections and the RAM bit pattern is fixed for most designs during product life cycle. Programmability in the LUT circuit is no longer needed and no longer valuable to the user. The user may convert the design to a lower cost hard-wire ROM circuit. The programmed LUT choices are mapped from RAM to ROM. RAM outputs at logic one are mapped to ROM wires connected to power. RAM outputs at logic zero are mapped to ROM wires connected to ground. This may be done with a single metal mask in lieu of all of the TFT layers. Such an elimination of processing layers reduces the cost of the ROM version. A first module with macro-LUT transistors does not change by this conversion. A third module may exist above the second module to complete interconnect for functionality of the end device. The third module also does not change with the

second module option. A timing characteristic comprising signal delay for LUT values to reach LUT output is not changed by the memory option. The propagation delays and critical path timing in the FPGA may be substantially identical between the two second module options. The TFT layers may allow a higher power supply voltage for the user to emulate performance at reduced pass-gate resistances. Such emulations may predict potential performance improvements for TFT pass-gates and hard-wired connected options. Duplicated ROM pattern may be done with a customized thru-hole mask. Customization may be done with a thru-hole and a metal mask or a plurality of thru-hole and metal masks. Hard wire pattern may also improve reliability and reduce defect density of the final product. The ROM pattern provides a cost economical final macro-LUT circuit to the user at a very low NRE cost. The total solution provides a programmable and customized solution to the user.

Implementations of the above aspect may further include one or more of the following. The programmable LUT circuit comprises a RAM element that can be selected from the group consisting of volatile or non volatile memory elements. The memory can be implemented using a TFT process technology that contains one or more of Fuses, Anti-fuses, DRAM, EPROM, EEPROM, Flash, Ferro-Electric, optical, magnetic, electro-chemical and SRAM elements. Configuration circuits may include thin film elements such as diodes, transistors, resistors and capacitors. The process implementation is possible with any memory technology where the programmable element is vertically integrated in a removable module. The manufacturing options include a conductive ROM pattern in lieu of memory circuits to control the logic in LUT circuits. Multiple memory bits exist to customize wire connections inside macro-LUTs,

inside a logic block and between logic blocks. Each RAM bit pattern has a corresponding unique ROM pattern to duplicate the same functionality.

The programmable LUT structures described constitutes fabricating a VLSI IC product. The IC product is re-programmable in its initial stage with turnkey conversion to a one mask customized ASIC. The IC has the end ASIC cost structure and initial FPGA re-programmability. The IC product offering occurs in two phases: the first phase is a generic FPGA that has re-programmability contained in a programmable LUT and programmable wire circuit, and a second phase is an ASIC that has the entire programmable module replaced by one or two customized hard-wire masks. Both FPGA version and turnkey custom ASIC has the same base die. No re-qualification is required by the conversion. The vertically integrated programmable module does not consume valuable Silicon real estate of a base die. Furthermore, the design and layout of these product families adhere to removable module concept: ensuring the functionality and timing of the product in its FPGA and ASIC canonicals. These IC products can replace existing PLD's, CPLD's, FPGA's, Gate Arrays, Structured ASIC's and Standard Cell ASIC's. An easy turnkey customization of an end ASIC from an original smaller cheaper and faster programmable structured array device would greatly enhance time to market, performance, product reliability and solution cost.

## **BRIEF DESCRIPTION OF THE DRAWINGS**

Fig-1A shows an exemplary MUX or LUT based logic element.

Fig-1B shows an exemplary programmable wire structure utilizing a logic element.

Fig-2A shows a truth table for a four variable function and the logic equation.

5 Fig-2B shows a 3-control-variable MUX realization of the function shown in Fig-2A.

Fig-2C shows a MUX input connection for a programmable version of MUX in Fig-2B.

Fig-2D shows an AND/OR gate realization of the function shown in Fig-2A.

Fig-2E shows a 4-input LUT realization of the function shown in Fig-2A.

Fig-3A shows an exemplary one input LUT (1LUT).

10 Fig-3B shows the symbol for 1LUT in Fig-3A that is used in rest of Fig-3.

Fig-3C – Fig-3F shows exemplary 2LUT, 3LUT, 4LUT and 5LUT respectively.

Fig-3G shows the number of 1LUTs needed to construct a K-LUT, where K is an integer from 1 to 7.

Fig-4 shows Silicon utilization efficiency with K-LUTs, extracted from Fig-10 in Ref-3.

15 Fig-5A shows an exemplary fuse link point to point connection.

Fig-5B shows an exemplary anti-fuse point to point connection.

Fig-5C shows an exemplary pass-gate point to point connection.

Fig-5D shows an exemplary floating-pass-gate point to point connection.

Fig-6A shows an exemplary configuration circuit for a 6T SRAM element.

20 Fig-6B shows an exemplary programmable pass-gate switch with SRAM memory.

Fig-7 shows an anti-fuse based configuration circuit.

Fig-8A shows a first embodiment of a floating gate configuration circuit.

Fig-8B shows a second embodiment of a floating gate configuration circuit.



Fig-9 shows a modular construction of a LUT circuit with removable TFT layers.

Fig-10.1-10.7 shows process cross-sections of one embodiment to integrate thin-film transistors into a logic process in accordance with the current invention.

Fig-11A shows a novel programmable 1-input LUT (1LUT).

5 Fig-11B shows the 1LUT in Fig-11A with a programmable MUX to select LUT values.

Fig-11C shows the 1LUT block diagram in Fig-11A with a configurable LUT value.

Fig-11D shows the 1LUT block diagram in Fig-11A with two configurable LUT values.

Fig-12A shows a second embodiment of a novel programmable 1LUT.

Fig-12B shows a third embodiment of a novel programmable 1LUT.

10 Fig-13A shows a fourth embodiment of a novel programmable 1LUT.

Fig-13B shows a fifth embodiment of a novel programmable 1LUT.

Fig-14 shows a novel programmable 2LUT macro-cell.

Fig-15 shows a novel programmable 3LUT macro-cell.

Fig-16A shows a first embodiment of a novel programmable 4LUT macro-cell.

15 Fig-16B shows a second embodiment of a novel programmable 4LUT macro-cell.

Fig-17A shows a first embodiment of a novel programmable 3LUT.

Fig-17B shows a second embodiment of a novel programmable 3LUT.

Fig-18A shows a truth table and logic equation of an example.

Fig-18B shows a 2LUT gate realization of the logic function in Fig-18A.

20 Fig-18C shows a 4LUT gate realization of the logic function in Fig-18B.

Fig-18D shows a programmable 4LUT gate realization of logic function in Fig-18B.

## DESCRIPTION

In the following detailed description of the invention, reference is made to the accompanying drawings which form a part hereof, and in which is shown, by way of illustration, specific embodiments in which the invention may be practiced. These  
5   embodiments are described in sufficient detail to enable those skilled in the art to practice the invention. Other embodiments may be utilized and structural, logical, and electrical changes may be made without departing from the scope of the present invention.

Definitions: The terms wafer and substrate used in the following description include any structure having an exposed surface with which to form the integrated circuit  
10   (IC) structure of the invention. The term substrate is understood to include semiconductor wafers. The term substrate is also used to refer to semiconductor structures during processing, and may include other layers that have been fabricated thereupon. Both wafer and substrate include doped and undoped semiconductors, epitaxial semiconductor layers supported by a base semiconductor or insulator, SOI material as well as other  
15   semiconductor structures well known to one skilled in the art. The term conductor is understood to include semiconductors, and the term insulator is defined to include any material that is less electrically conductive than the materials referred to as conductors.

The term module layer includes a structure that is fabricated using a series of predetermined process steps. The boundary of the structure is defined by a first step, one  
20   or more intermediate steps, and a final step. The resulting structure is formed on a substrate.

The term pass-gate refers to a structure that can pass a signal when on, and blocks signal passage when off. A pass-gate connects two points when on, and disconnects two

points when off. A pass-gate can be a floating-gate transistor, an NMOS transistor, a PMOS transistor or a CMOS transistor pair. The gate electrode of pass-gate determines the state of the connection. A CMOS pass-gate requires complementary signals coupled to NMOS and PMOS gate electrodes. A control logic signal is connected to gate  
5 electrode of a pass-gate for programmable logic.

The term configuration circuit includes one or more configurable elements and connections that can be programmed for controlling one or more circuit blocks in accordance with a predetermined user-desired functionality. The configuration circuit includes the memory element and the access circuitry, herewith called memory circuitry,  
10 to modify said memory element. Configuration circuit does not include the logic pass-gate controlled by said memory element. In one embodiment, the configuration circuit includes a plurality of RAM circuits to store instructions to configure an FPGA. In another embodiment, the configuration circuit includes a first selectable configuration where a plurality of RAM circuits is formed to store instructions to control one or more  
15 circuit blocks. The configuration circuits include a second selectable configuration with a predetermined ROM conductive pattern formed in lieu of the RAM circuit to control substantially the same circuit blocks. The memory circuit includes elements such as diode, transistor, resistor, capacitor, metal link, wires, among others. The memory circuit also includes thin film elements. In yet another embodiment, the configuration circuits  
20 include a predetermined conductive pattern, contact, via, resistor, capacitor or other suitable circuits formed in lieu of the memory circuit to control substantially the same circuit blocks.

The term "horizontal" as used in this application is defined as a plane parallel to the conventional plane or surface of a wafer or substrate, regardless of the orientation of the wafer or substrate. The term "vertical" refers to a direction perpendicular to the horizontal direction as defined above. Prepositions, such as "on", "side", "higher",  
5 "lower", "over" and "under" are defined with respect to the conventional plane or surface being on the top surface of the wafer or substrate, regardless of the orientation of the wafer or substrate. The following detailed description is, therefore, not to be taken in a limiting sense.

The term K-LUT refers to a look up table comprising K inputs. Such a LUT  
10 comprises  $2^K$  LUT values, and at least one output. For a given combination of K-input values, a LUT value is received at said at least one LUT output. The term LUT tree and LUT cone refers to construction of a LUT, where there is a gradual decrease in the number of LUTs in each stage. A first of the K-inputs is common to all the LUTs in a first stage, a second of the K-inputs is common to all the LUTs in a second stage and so  
15 on until the last LUT stage is reached in a hard wired K-LUT tree.

Programmable LUTs use point to point connections that utilize programmable pass-gate logic as shown in Fig-6A and Fig-6B. Multiple inputs (node A) can be connected to multiple outputs (node B) with a plurality of pass-gate logic elements. The SRAM base connection shown in Fig-6 may have pass-gate 610 as a PMOS or an NMOS  
20 transistor. NMOS is preferred due to its higher conduction. The voltage  $S_0$  on NMOS transistor 610 gate electrode determines an ON or OFF connection. That logic level is generated by a configuration circuit 650 coupled to the gate of NMOS transistor 610. The pass-gate logic connection requires the configuration circuitry to generate signal  $S_0$  with

sufficient voltage levels to ensure off and on conditions. For an NMOS pass-gate,  $S_0$  having a logic level one completes the point to point connection, while a logic level zero keeps them disconnected. In addition to using only an NMOS gate, a PMOS gate could also be used in parallel to make the connection. The configuration circuit 650 needs to  
5 then provide complementary outputs ( $S_0$  and  $S_0'$ ) to drive NMOS and PMOS gates in the connection. Configuration circuit 650 contains a memory element. Most CMOS SRAM memory delivers complementary outputs. This memory element can be configured by the user to select the polarity of  $S_0$ , thereby selecting the status of the connection. The memory element can be volatile or non-volatile. In volatile memory, it could be DRAM,  
10 SRAM, Optical or any other type of a memory device that can output a valid signal  $S_0$ . In non-volatile memory it could be fuse, anti-fuse, EPROM, EEPROM, Flash, Ferro-Electric, Magnetic or any other kind of memory device that can output a valid signal  $S_0$ . The output  $S_0$  can be a direct output coupled to the memory element, or a derived output in the configuration circuitry. An inverter can be used to restore  $S_0$  signal level to full rail  
15 voltage levels. The SRAM in configuration circuit 650 can be operated at an elevated  $V_{cc}$  level to output an elevated  $S_0$  voltage level. This is especially feasible when the SRAM is built in a separate TFT module. Other configuration circuits to generate a valid  $S_0$  signal are discussed next.

An anti-fuse based configuration circuit to use with this invention is shown next  
20 in Fig-7. Configuration circuit 650 in Fig-6B can be replaced with the anti-fuse circuit shown in Fig-7. In Fig-7, output level  $S_0$  is generated from node X which is coupled to signals VA and VB via two anti-fuses 750 and 760 respectively. Node X is connected to a programming access transistor 770 controlled by gate signal GA and drain signal BL. A

very high programming voltage is needed to blow the anti-fuse capacitor. This programming voltage level is determined by the anti-fuse properties, including the dielectric thickness. Asserting signal VA very high, VB low (typically ground), BL low and GA high ( $V_{cc}$  to pass the ground signal) provides a current path from VA to BL through the on transistor 770. A high voltage is applied across anti-fuse 750 to pop the dielectric and short the terminals. Similarly anti-fuse 760 can be programmed by selecting VA low, VB very high, BL low and GA high. Only one of the two anti-fuses is blown to form a short. When the programming is done, BL and GA are returned to zero, isolating node X from the programming path.  $VA=V_{ss}$  (ground) and  $VB=V_{cc}$  (power, or elevated  $V_{cc}$ ) is applied to the two signal lines. Depending on the blown fuse, signal  $S_0$  will generate a logic low or a logic high signal. This is a one time programmable memory device. Node X will be always connected to VA or VB by the blown fuse regardless of the device power status. Signals GA and BL are constructed orthogonally to facilitate row and column based decoding to construct these memory elements in an array.

Fig-8 shows two EEPROM non-volatile configuration circuits that can be used in this invention. Configuration circuit 650 in Fig-6B can be replaced with either of two EEPROM circuit shown in Fig-8A and Fig-8B. In Fig-8A, node 840 is a floating gate. This is usually a poly-Silicon film isolated by an insulator all around. It is coupled to the source end of programming transistor 820 via a tunneling diode 830. The tunneling diode is a thin dielectric capacitor between floating poly and substrate Silicon with high doping on either side. When a large programming (or erase) voltage  $V_{pp}$  is applied across the thin dielectric, a Fowler-Nordheim tunneling current flows through the oxide. The tunneling electrons move from electrical negative to electrical positive voltage. Choosing

the polarity of the applied voltage across the tunneling dielectric, the direction of electron flow can be reversed. Multiple programming and erase cycles are possible for these memory elements. As the tunneling currents are small, the high programming voltage ( $V_{pp}$ ) can be generated on chip, and the programming and erasure can be done while the

5 chip is in a system. It is hence called in system programmable (ISP). An oxide or dielectric capacitor 810 couples the floating gate (FG) 840 to a control gate (CG). The control gate CG can be a heavily doped Silicon substrate plate or a second poly-Silicon plate above the floating poly. The dielectric can be oxide, nitride, ONO or any other insulating material. A voltage applied to CG will be capacitively coupled to FG node 840.

10 The coupling ratio is designed such that 60-80 percent of CG voltage will be coupled to FG node 840. To program this memory element, a negative charge must be trapped on the FG 840. This is done by applying positive  $V_{pp}$  voltage on CG, ground voltage on PL and a sufficiently high ( $V_{cc}$ ) voltage on RL. CG couples a high positive voltage onto FG 840 creating a high voltage drop across diode 830. Electrons move to the FG 840 to

15 reduce this electric field. When the memory device is returned to normal voltages, a net negative voltage remains trapped on the FG 840. To erase the memory element, the electrons must be removed from the floating gate. This can be done by UV light, but an electrical method is more easily adapted. The CG is grounded, a very high voltage ( $V_{pp}$  + more to prevent a threshold voltage drop across 820) is applied to RL, and a very high

20 voltage ( $V_{pp}$ ) is applied to PL. Now a low voltage is coupled to FG with a very high positive voltage on the source side of device 820. Diode 830 tunneling removes electrons from FG. This removal continues beyond a charge neutral state for the isolated FG. When the memory device is returned to normal voltages, a net positive voltage remains trapped

on the FG 840. Under normal operation RL is grounded to isolate the memory element from the programming path, and PL is grounded. A positive intermediate voltage  $V_{cg}$  is applied to CG terminal. FG voltage is denoted  $S_0$ . Under CG bias,  $S_0$  signal levels are designed to activate pass-gate logic correctly. Configuration circuit in Fig-8B is only  
5 different to that in Fig-8A by the capacitor 851 used to induce  $S_0$  voltage. This is useful when  $S_0$  output is applied to leaky pass-gates, or low level leakage nodes. As gate oxide thicknesses reach below 50 angstroms, the pass-gates leak due to direct tunneling.

These configuration circuits, and similarly constructed other configuration circuits, can be used in programmable logic devices. Those with ordinary skill in the art  
10 may recognize other methods for constructing configuration circuits to generate a valid  $S_0$  output. The pass-gate logic element is not affected by the choice of the configuration circuit.

SRAM memory technology has the advantage of not requiring a high voltage to configure memory. The SRAM based switch shown in Fig-6B containing the SRAM  
15 memory circuit shown in Fig-6A utilizes 6 extra configuration transistors, discounting the pass-gate 610, to provide the programmability. That is a significant overhead compared to application specific and hard-wired gate array circuits where the point to point connection can be directly made with metal. Similarly other programmable memory elements capable of configuring pass-gate logic also carry a high Silicon foot print. A  
20 cheaper method of constructing a vertically integrated SRAM cell is described in incorporated by reference Application Serial No. 10/413,810. In a preferred embodiment, the configuration circuit is built on thin-film semiconductor layers located vertically above the logic circuits. The SRAM memory element, a thin-film transistor (TFT) CMOS



latch as shown in Fig-6A, comprises two lower performance back to back inverters formed on two semiconductor thin film layers, substantially different from a first semiconductor single crystal substrate layer and a gate poly layer used for logic transistor construction. This latch is stacked above the logic circuits for slow memory applications with no penalty on Silicon area and cost. This latch is adapted to receive power and ground voltages in addition to configuration signals. The two programming access transistors for the TFT latch are also formed on thin-film layers. Thus in Fig-6B, all six configuration transistors shown in 650 are constructed in TFT layers, vertically above the pass transistor 610. Transistor 610 is in the conducting path of the connection and needs to be a high performance single crystal Silicon transistor. This vertical integration makes it economically feasible to add an SRAM based configuration circuit at a very small cost overhead to create a programmable solution. Such vertical integration can be extended to all other memory elements that can be vertically integrated above logic circuits.

A new kind of a programmable logic device utilizing thin-film transistor configurable circuits is disclosed in incorporated by reference Application Serial No. 10/267483, Application Serial No. 10/267484 and Application Serial No. 10/267511. The disclosures describe a programmable logic device and an application specific device fabrication from the same base Silicon die. The PLD is fabricated with a programmable RAM module, while the ASIC is fabricated with a conductive ROM pattern in lieu of the RAM. Both RAM module and ROM module provide identical control of logic circuits. For each set of RAM bit patterns, there is a unique ROM pattern to achieve the same logic functionality. The vertical integration of the configuration circuit leads to a significant cost reduction for the PLD, and the elimination of TFT memory for the ASIC

allows an additional cost reduction for the user. The TFT vertical memory integration scheme is briefly described next.

Fig-9 shows an implementation of vertically integrated circuits, where the configuration memory element is located above logic. The memory element can be any one of fuse links, anti-fuse capacitors, SRAM cells, DRAM cells, metal optional links, EPROM cells, EEPROM cells, flash cells, ferro-electric elements, electro-chemical elements, optical elements and magnetic elements that lend to this implementation. SRAM memory is used herein to illustrate the scheme and is not to be taken in a limiting sense. First, Silicon transistors 950 are deposited on a substrate. A module layer of removable SRAM cells 952 are positioned above the Silicon transistors 950, and a module layer of interconnect wiring or routing circuit 954 is formed above the removable memory cells 952. To allow this replacement, the design adheres to a hierarchical layout structure. As shown in Fig-9, the SRAM cell module is sandwiched between the single crystal device layers below and the metal layers above electrically connecting to both. It also provides through connections "A" for the lower device layers to upper metal layers. The SRAM module contains no switching electrical signal routing inside the module. All such routing is in the layers above and below. Most of the programmable element configuration signals run inside the module. Upper layer connections to SRAM module "C" are minimized to Power, Ground and high drive data wires. Connections "B" between SRAM module and single crystal module only contain logic level signals and replaced later by Vcc and Vss wires. Most of the replaceable programmable elements and its configuration wiring is in the "replaceable module" while all the devices and wiring for the end ASIC is outside the "replaceable module". In other embodiments, the

replaceable module could exist between two metal layers or as the top most module layer satisfying the same device and routing constraints. This description is equally applicable to any other configuration memory element, and not limited to SRAM cells.

Fabrication of the IC also follows a modularized device formation. Formation of transistors 950 and routing 954 is by utilizing a standard logic process flow used in the ASIC fabrication. Extra processing steps used for memory element 952 formation are inserted into the logic flow after circuit layer 950 is constructed. A full disclosure of the vertical integration of the TFT module using extra masks and extra processing is in the incorporated by reference applications listed above.

During the ROM customization, the base die and the data in those remaining mask layers do not change making the logistics associated with chip manufacture simple. Removal of the SRAM module provides a low cost standard logic process for the final ASIC construction with the added benefit of a smaller die size. The design timing is unaffected by this migration as lateral metal routing and Silicon transistors are untouched. Software verification and the original FPGA design methodology provide a guaranteed final ASIC solution to the user. A full disclosure of the ASIC migration from the original FPGA is in the incorporated by reference applications discussed above.

In Fig-9, the third module layer is formed substantially above the first and second module layers, wherein interconnect and routing signals are formed to connect the circuit blocks within the first and second module layers. Alternatively, the third module layer can be formed substantially below the first and second module layer with interconnect and routing signals formed to connect the circuit blocks within the first and second module layers. Alternatively, the third and fourth module layers positioned above and

below the second module layer respectively, wherein the third and fourth module layers provide interconnect and routing signals to connect the circuit blocks within the first and second module layers.

In yet another embodiment of a programmable multi-dimensional semiconductor device, a first module layer is fabricated having a plurality of circuit blocks formed on a first plane. The programmable multi-dimensional semiconductor device also includes a second module layer formed on a second plane. A plurality of configuration circuits is then formed in the second plane to store instructions to control a portion of the circuit blocks.

The fabrication of thin-film transistors to construct configuration circuits is discussed next. A full disclosure is provided in incorporated by reference Application Serial Number 10/413809. The following terms used herein are acronyms associated with certain manufacturing processes. The acronyms and their abbreviations are as follows:

$V_T$  Threshold voltage

LDN Lightly doped NMOS drain

LDP Lightly doped PMOS drain

LDD Lightly doped drain

RTA Rapid thermal annealing

Ni Nickel

Co Cobalt

Ti Titanium

TiN Titanium-Nitride

W Tungsten

	S	Source
	D	Drain
	G	Gate
	ILD	Inter layer dielectric
5	C1	Contact-1
	M1	Metal-1
	P1	Poly-1
	P-	Positive light dopant (Boron species, $\text{BF}_2$ )
	N-	Negative light dopant (Phosphorous, Arsenic)
10	P+	Positive high dopant (Boron species, $\text{BF}_2$ )
	N+	Negative high dopant (Phosphorous, Arsenic)
	Gox	Gate oxide
	C2	Contact-2
	LPCVD	Low pressure chemical vapor deposition
15	CVD	Chemical vapor deposition
	ONO	Oxide-nitride-oxide
	LTO	Low temperature oxide

A logic process is used to fabricate CMOS devices on a substrate layer for the fabrication of logic circuits. These CMOS devices may be used to build AND gates, OR gates, inverters, adders, multipliers, memory and pass-gate based logic functions in an integrated circuit. A CMOSFET TFT module layer or a Complementary gated FET (CGated-FET) TFT module layer may be inserted to a logic process at a first contact mask to build a second set of TFT MOSFET or Gated-FET devices. Configuration

circuitry including RAM elements is build with these second set of transistors. An exemplary logic process may include one or more following steps:

P-type substrate starting wafer

Shallow Trench isolation: Trench Etch, Trench Fill and CMP

5 Sacrificial oxide deposition

PMOS  $V_T$  mask & implant

NMOS  $V_T$  mask & implant

Pwell implant mask and implant through field

Nwell implant mask and implant through field

10 Dopant activation and anneal

Sacrificial oxide etch

Gate oxidation / Dual gate oxide option

Gate poly (GP) deposition

GP mask & etch

15 LDN mask & implant

LDP mask & implant

Spacer oxide deposition & spacer etch

N+ mask and NMOS N+ G, S, D implant

P+ mask and PMOS P+ G, S, D implant

20 Co deposition

RTA anneal - Co salicidation (S/D/G regions & interconnect)

Unreacted Co etch

ILD oxide deposition & CMP

Fig-10 shows an exemplary process for fabricating a thin film MOSFET latch in a second module layer. In one embodiment the process in Fig-10 forms the latch in a layer substantially above the substrate layer. The processing sequence in Fig-10.1 through Fig-10.7 describes the physical construction of a MOSFET device for storage circuits 650 shown in Fig-6B. The process of Fig-10 includes adding one or more following steps to the logic process after ILD oxide deposition & CMP step in the logic process.

C1 mask & etch

W-Silicide plug fill & CMP

~250A poly P1 (amorphous poly-1) deposition

10 P1 mask & etch

Blanket Vtn P- implant (NMOS Vt)

Vtp mask & N- implant (PMOS Vt)

TFT Gox (70A PECVD) deposition

400A P2 (amorphous poly-2) deposition

15 P2 mask & etch

Blanket LDN NMOS N- tip implant

LDP mask and PMOS P- tip implant

Spacer LTO deposition

Spacer LTO etch to form spacers & expose P1

20 Blanket N+ implant (NMOS G/S/D & interconnect)

P+ mask & implant (PMOS G/S/D & interconnect).

Ni deposition

RTA salicidation and poly re-crystallization (G/S/D regions & interconnect)

Dopant activation anneal

Excess Ni etch

ILD oxide deposition & CMP

C2 mask & etch

5 W plug formation & CMP

M1 deposition and back end metallization

The TFT process technology consists of creating NMOS & PMOS poly-Silicon transistors. In the embodiment in Fig-10, the module insertion is after the substrate device  
10 gate-poly etch and ILD film deposition. In other embodiments the insertion point may be after M1 and ILD deposition, prior to V1 mask, or between two metal definition steps.

After gate poly of regular transistors are patterned and etched, the poly is salicided using Cobalt & RTA sequences. Then the ILD is deposited, and polished by CMP techniques to a desired thickness. In the shown embodiment, the contact mask is  
15 split into two levels. The first C1 mask contains all contacts that connect TFT latch outputs to substrate transistor pass-gates. This C1 mask is used to open and etch contacts in the ILD film. Ti/TiN glue layer followed by W-Six plugs, W plugs or Si plugs may be used to fill the plugs, then CMP polished to leave the fill material only in the contact holes. The choice of fill material is based on the thermal requirements of the TFT  
20 module. In another embodiment, Ni is introduced into C1 to facilitate crystallization of the poly Silicon deposited over the contacts. This Ni may be introduced as a thin layer after the Ti/TiN glue layer is deposited, or after W is deposited just to fill the center of the contact hole.



Then, a desired thickness of first P1 poly, amorphous or crystalline, is deposited by LPCVD as shown in Fig-10.1. The P1 thickness is between 50A and 1000A, and preferably 250A. This poly layer P1 is used for the channel, source, and drain regions for both NMOS and PMOS TFT's. It is patterned and etched to form the transistor body regions. In other embodiments, P1 is used for contact pedestals. NMOS transistors are blanket implanted with P- doping, while the PMOS transistor regions are mask selected and implanted with N- doping. This is shown in Fig-10.2. The implant doses and P1 thickness are optimized to get the required threshold voltages for PMOS & NMOS devices under fully depleted transistor operation, and maximize on/off device current ratio. The pedestals implant type is irrelevant at this point. In another embodiment, the  $V_T$  implantation is done with a mask P- implant followed by masked N- implant. First doping can also be done in-situ during poly deposition or by blanket implant after poly is deposited.

Patterned and implanted P1 may be subjected to dopant activation and crystallization. In one embodiment, an RTA cycle with Ni as seed in C1 is used to activate & crystallize the poly before or after it is patterned to near single crystal form. In a second embodiment, the gate dielectric is deposited, and buried contact mask is used to etch areas where P1 contacts P2 layer. Then, Ni is deposited and salicided with RTA cycle. All of the P1 in contact with Ni is salicided, while the rest poly is crystallized to near single crystal form. Then the un-reacted Ni is etched away. In a third embodiment, amorphous poly is crystallized prior to P1 patterning with an oxide cap, metal seed mask, Ni deposition and MILC (Metal-Induced-Lateral-Crystallization).

Then the TFT gate dielectric layer is deposited followed by P2 layer deposition. The dielectric is deposited by PECVD techniques to a desired thickness in the 30-200A range, desirably 70A thick. The gate may be grown thermally by using RTA. This gate material could be an oxide, nitride, oxynitride, ONO structure, or any other dielectric material combinations used as gate dielectric. The dielectric thickness is determined by the voltage level of the process. At this point an optional buried contact mask (BC) may be used to open selected P1 contact regions, etch the dielectric and expose P1 layer. BC could be used on P1 pedestals to form P1/P2 stacks over C1. In the P1 salicided embodiment using Ni, the dielectric deposition and buried contact etch occur before the crystallization. In the preferred embodiment, no BC is used.

Then second poly P2 layer, 100A to 2000A thick, preferably 400A is deposited as amorphous or crystalline poly-Silicon by LPCVD as shown in Fig-10.3. P2 layer is defined into NMOS & PMOS gate regions intersecting the P1 layer body regions, C1 pedestals if needed, and local interconnect lines and then etched. The P2 layer etching is continued until the dielectric oxide is exposed over P1 areas uncovered by P2 (source, drain, P1 resistors). The source & drain P1 regions orthogonal to P2 gate regions are now self aligned to P2 gate edges. The S/D P2 regions may contact P1 via buried contacts. NMOS devices are blanket implanted with LDN N- dopant. Then PMOS devices are mask selected and implanted with LDP P- dopant as shown in Fig-10.4. The implant energy ensures full dopant penetration through the residual oxide into the S/D regions adjacent to P2 layers.

A spacer oxide is deposited over the LDD implanted P2 using LTO or PECVD techniques. The oxide is etched to form spacers. The spacer etch leaves a residual oxide

over P1 in a first embodiment, and completely removes oxide over exposed P1 in a second embodiment. The latter allows for P1 salicidation at a subsequent step. Then NMOS devices & N+ poly interconnects are blanket implanted with N+. The implant energy ensures full or partial dopant penetration into the 100A residual oxide in the S/D regions adjacent to P2 layers. This doping gets to gate, drain & source of all NMOS devices and N+ interconnects. The P+ mask is used to select PMOS devices and P+ interconnect, and implanted with P+ dopant as shown in Fig-10.5. PMOS gate, drain & source regions receive the P+ dopant. This N+/P+ implants can be done with N+ mask followed by P+ mask. The  $V_T$  implanted P1 regions are now completely covered by P2 layer and spacer regions, and form channel regions of NMOS & PMOS transistors.

After the P+/N+ implants, Nickel is deposited over P2 and salicided to form a low resistive refractory metal on exposed poly by RTA. Un-reacted Ni is etched as shown in Fig-10.6. This 100A-500A thick Ni-Salicide connects the opposite doped poly-2 regions together providing low resistive poly wires for data. In one embodiment, the residual gate dielectric left after the spacer prevents P1 layer salicidation. In a second embodiment, as the residual oxide is removed over exposed P1 after spacer-etch, P1 is salicided. The thickness of Ni deposition may be used to control full or partial salicidation of P1 regions. Fully salicided S/D regions up to spacer edge facilitate high drive current due to lower source and drain resistances.

An LTO film is deposited over P2 layer, and polished flat with CMP. A second contact mask C2 is used to open contacts into the TFT P2 and P1 regions in addition to all other contacts to substrate transistors. In the shown embodiment, C1 contacts connecting latch outputs to substrate transistor gates require no C2 contacts. Contact

plugs are filled with tungsten, CMP polished, and connected by metal as done in standard contact metallization of IC's as shown in Fig-10.7.

A TFT process sequence similar to that shown in Fig-10 can be used to build complementary Gated-FET thin film devices. Compared with CMOS devices, these are bulk conducting devices and work on the principles of JFETs. A full disclosure of these devices is provided in incorporated by reference Application Serial No. 10/413,808. The process steps facilitate the device doping differences between MOSFET and Gated-FET devices, and simultaneous formation of complementary Gated-FET TFT devices. A detailed description for this process was provided when describing Fig-10 earlier and is not repeated. An exemplary CGated-FET process sequence may use one or more of the following steps:

C1 mask & etch

W-Silicide plug fill & CMP (optional Ni seed in W-plug)

~300Å poly P1 (amorphous poly-1) deposition

15 Optional poly crystallization

P1 mask & etch

Blanket V<sub>tn</sub> N- implant (Gated-NFET V<sub>T</sub>)

V<sub>tp</sub> mask & P- implant (Gated-PFET V<sub>T</sub>)

TFT Gox (70Å PECVD) deposition

20 500Å P2 (amorphous poly-2) deposition

Blanket P+ implant (Gated-NFET gate & interconnect)

N+ mask & implant (Gated-PFET gate & interconnect)

P2 mask & etch

Blanket LDN Gated-NFET N tip implant

LDP mask and Gated-PFET P tip implant

Spacer LTO deposition

Spacer LTO etch to form spacers & expose P1

5 Ni deposition

RTA salicidation and poly re-crystallization (exposed P1 and P2)

Fully salicidation of exposed P1 S/D regions

Dopant activation anneal

Excess Ni etch

10 ILD oxide deposition & CMP

C2 mask & etch

W plug formation & CMP

M1 deposition and back end metallization

15 As the discussions demonstrate, memory controlled pass transistor logic elements provide a powerful tool to make switches. The ensuing high cost of memory can be drastically reduced by the 3-dimensional integration of configuration elements and the replaceable modularity concept for said memory. These advances allow designing a LUT based macrocell with more programmable bits to overcome the deficiencies associated

20 with logic fitting in large LUT sizes. In one aspect, a cheaper memory element allows use of more memory for programmability. That enhances the ability to build large logic blocks utilizing multiple LUTs (i.e. course-grain advantage) while maintaining smaller logic element type logic fitting (i.e. fine-grain advantage). Furthermore larger grains need

less connectivity: neighboring cells and far-away cells. That further simplifies the interconnect structure. Larger grains benefit by larger LUT sizes, or a larger number of bigger LUTs in a logic block. In a second aspect cheaper memory allows LUT partitioning that can efficiently utilize Silicon by fitting large and small logic pieces into a single large LUT. Such LUTs can improve Silicon utilization compared to Fig-4. A new programmable LUT macrocell circuit utilizing the manufacturing methods shown so far is discussed next. Larger LUT integration is discussed by Wittig et al. US 6208163, Agrawal et al. US 2002/0186044, Sueyoshi et al. US 2003/0001615 and Pugh et al. US 2003/0085733. They do not show the need, a method and the value in using programmable bits to provide multiple smaller LUT partitioning inside a single larger LUT for FPGA designs.

A one input LUT (1LUT) according to current teaching is shown in Fig-11A. The LUT is comprised of input A driving pass-gate 1101. Input compliment A' drives pass-gate 1102. Cross-circled elements 1111, 1112 & 1113 represent memory bits in a configurable memory circuit. An SRAM based memory circuit described earlier is shown in Fig-6. Such a memory circuit provides complimentary outputs  $S_0$  &  $S_0'$  to control on-off behavior of pass-gates 1101-1106. The LUT values are selected by programmable bit such as 1111 in one of two configurations. When the memory bit is programmed to a logic one, the bit 1111 outputs a logic one  $S_0$  on the right hand side branch and logic zero  $S_0'$  on the left hand branch. When the memory bit is programmed to a logic zero, the bit 1111 outputs a logic zero  $S_0$  on the right hand side branch and logic one  $S_0'$  on the left hand branch. This allows selecting  $I_1$ ,  $I_2$  pair as LUT values by setting memory bit 1111 to zero, or selecting values stored in register 1112, 1113 pair as LUT values by setting

memory bit 1111 to one. The inputs  $I_1$  and  $I_2$  are also driven by buffers that are not shown in Fig-11A. Memory bits 1111, 1112 & 1113 are constructed in a thin-film module and are vertically integrated. TFT SRAM 1112 and 1113 drive inverters constructed in substrate Silicon or pass-gates coupling  $V_{cc}$  &  $V_{ss}$  to provide necessary LUT value drive currents. All TFT memory circuits allow the user to change stored data as desired. The configuration circuits including memory is constructed over the pass-gate logic circuits and consumes no Silicon area and cost. When selected, the registers 1112 & 1113 can be independently set to logic states one or zero by the user, and becomes identical to the 1LUT shown in Fig-3A. Once the desired memory pattern is identified by the user, TFT elements 1111, 1112 & 1113 can be replaced by hard-wires connected to  $V_{cc}$  or  $V_{ss}$  to achieve identical logic functionality. As the timing path is restricted to signal propagation in wires and pass-gates, there is no change in timing with this conversion. As the fabrication process is simplified by eliminating TFT memory processing, the end product is cheaper to fabricate and more reliable for the user.

Two Embodiments of block diagrams of the LUT shown in Fig-11A are shown in Fig-11C and Fig-11D. Referring to Fig-11C, a programmable look up table (LUT) circuit 1138 for an integrated circuit, comprises: one or more secondary inputs 1132; and one or more configurable logic states 1134; and two or more LUT values 1135, 1136; and a programmable means 1133 to select a LUT value from a secondary input 1132 or a configurable logic state 1134. Referring to Fig-11D, the circuit 1148 further comprises: a LUT output 1147; and  $M$  primary inputs such as 1141, where  $M$  is an integer value greater than or equal to one, each said  $M$  inputs received in true and compliment logic levels; and  $2^M$  LUT values such as 1145 & 1146, each said LUT values comprising a

configurable logic state or a secondary input, wherein any given combination of said M primary input signal levels couples one of said LUT values to said LUT output.

An equivalent MUX representation for Fig-11A is shown in Fig-11B. The LUT values are chosen from two 3-input MUXs 1151 and 1152 with 3 programmable bits, wherein the gate construction is as in Fig-11A, and the block diagram is as in Fig-11D.

A second embodiment of a programmable 1LUT according to this teaching is shown in Fig-12A. This 1LUT utilizes 4-programmable memory bits 1211, 1212, 1213 and 1214, and otherwise identical to 1LUT in Fig-11A. Having 4 programmable bits allows the user to select the upper half of 1LUT independent of the lower half. For example, bit 1211 can be configured to select  $I_1$  as a LUT value for A input, and bit 1214 can be configured to select register 1213 as the LUT value for A' input. This flexibility in a LUT macrocell is extremely useful to reduce Silicon wastage as will be shown later.

Another embodiment of the programmable macro-cell according to these teachings utilizing 4-programmable bits is shown in Fig-12B. This has two 4:1 MUXs 1351 and 1352 that are configured by 2 bits each for each LUT value. Each 4:1 MUX is identical to the MUX shown in Fig-2C. LUT value for input A is programmed from  $I_1$ ,  $I_2$ , 0 & 1, while LUT value for input A' is programmed from  $I_3$ ,  $I_4$ , 0 & 1. This 1LUT macro-cell allows the user to select which inputs needs to couple from previous to next LUT stage. When  $I_1=I_3=B$  and  $I_2=I_4=B'$ , Fig-12B becomes a 2-input LUT. Memory circuits for Fig-12 are also constructed in TFT layers to occupy no extra Silicon area.

A third embodiment of a programmable 1LUT according to this teaching is shown in Fig-13A. This 1LUT also utilizes 4-programmable memory bits 1311, 1312, 1313 and 1314, but provides an option for inputs  $I_1$  and  $I_2$  to by-pass the 1LUT. Otherwise, Fig-



13A is identical to 1LUT in Fig-12A. Bit 1311 polarity controls both logic state 1312 selection and input  $I_1$  by-pass. When LUT values are chosen to be logic states from 1312 & 1313, the inputs 1321 & 1322 are by-passed to registers not shown in the Fig-13A. The circuit shown in Fig-13A has a programmable method 1311 further comprising a means  
5 of providing said secondary input 1321 as an output when said configurable logic state 1312 is selected as a LUT value. Secondary input 1312 is provided as an output via the by-pass pass-gate 1308. Having 4 programmable bits allows the user to select the upper half of 1LUT independent of the lower half. For example, bit 1311 can be configured to select  $I_1$  as a LUT value for A input and disable  $I_1$  by-pass pass-gate 1308. Bit 1314 can  
10 be configured to select register 1313 output as the LUT value for A' input and shunt  $I_2$  input to an output register through pass-gate 1303. This flexibility in a LUT macrocell is also useful to reduce Silicon wastage as will be shown later. Yet another embodiment of the programmable macro-cell according to these teachings utilizing 6-programmable bits is shown in Fig-13B. This has two 8:1 MUXs 1351 and 1352 that are configured by 3 bits  
15 each. Each 8:1 MUX is a conventional MUX similar to the 4:1 MUX shown in Fig-2C. Upper half of 1LUT and lower half of 1LUT are independently programmed to one of eight choices for that LUT value. Apart from 0 and 1, the remaining 6 LUT value choices need not be identical. This LUT macro-cell allows the user to select multiple inputs in a LUT structure to perform a logic function of two variables. Memory circuits for Fig-13  
20 are constructed in TFT layers.

A 2-input LUT construction from programmable 1LUTs is shown in Fig-14. The 2LUT has 4 LUT values in registers 1421, 1422, 1423 and 1424. These LUT values are controlled by common input B on pass-gates 1401, 1402, 1403 and 1404. The outputs

from this first stage are fed to a programmable 1LUT similar to the one discussed in Fig-13A. Four programmable registers 1425, 1426, 1427 and 1428 control the second stage 1LUT providing the capability of combining the 2 LUTs or using them independently.

A 3-input LUT (3LUT) according to present invention is shown in Fig-15. Two conventional 2LUTs 1501 and 1502 are fed to a programmable 1LUT discussed in Fig-13A. This LUT macrocell can be configured to perform two independent 2LUT functions and one 1LUT function. The 2LUT outputs can by-pass the 1LUT and feed registers not shown in Fig-15. LUT macrocell can also perform one 3LUT function when C & E are made common and B & D are made common. In addition, the LUT macrocell can also perform a 3LUT (when the 3LUT function has half of the truth table entries as zero or one) plus a 2LUT. It can also perform some 4-input and 5-input variable functions. These divisions in logic allow improved logic fitting into LUT macrocells.

A 4-input LUT (4LUT) according to present invention is shown in Fig-16A and Fig-16B. In Fig-16A, four conventional 2LUTs 1601-1604 are fed to a programmable 2LUT 1605. The 2LUT 1605 is constructed with 2 programmable 1LUTs discussed in Fig-13A. This LUT macrocell can be configured to perform a wide variety of logic functions. It can perform five independent 2LUT functions, and all 2LUT outputs can be fed to registers (not shown). This is done by programming 2LUT 1605 to full independent mode by selecting all configurable states (such as 1613 & 1614) as LUT values. It can also perform one 4LUT function when first stage inputs (D, F, H, K) are made common and second stage inputs (C, E, G, J) are made common. There may be programmable switches to make these common inputs. When the 4LUT function has rows or columns in the truth table entries as zero or one, a LUT value is chosen in 2LUT

1605 to save a full 2LUT in a prior stage. Hence the LUT macrocell can also performs a 4LUT plus one or more 2LUTs to enhance logic density. It can also perform some 5-input, 6-input, up to 10-input variable functions. The LUT inputs are selected from a group of external inputs by programmable MUXs not shown in the diagram. These divisions in logic allow improved logic fitting into LUT macrocell based architectures. Compared to percentage logic overhead for 1LUT 1503 in Fig-15, the percentage overhead required for the added flexibility in 2LUT 1605 is lower in Fig-16A.

Referring to Fig-16A, A programmable look up table circuit 1605 for an integrated circuit, comprises: M primary inputs (such as A & B), wherein M is an integer value greater than or equal to one, and each said M inputs received in true and compliment logic levels; and  $2^M$  secondary inputs (such as 1611, 1612); and  $2^M$  configurable logic states (such as 1613, 1614), each said state comprising a logic zero and a logic one; and  $2^M$  LUT values; and a programmable means to select each of said LUT values from a secondary input (such as 1611) or a configurable logic state (such as 1613). In circuit 1605, each of said secondary inputs (such as 1611) is further comprised of an output of a previous K-LUT circuit (such as 1601), said K-LUT circuit comprising: a LUT output (same as 1611); and K inputs (such as C & D), wherein K is an integer value greater than or equal to one, and each said K inputs received in true and compliment logic levels; and  $2^K$  LUT values (such as crossed-circle latch outputs in 1601), each said LUT values comprising two configurable logic states.

Referring to Fig-16A, a larger N-LUT is constructed with smaller K-LUTs (such as 1601-1605). Each smaller K-LUT is further constructed as one of: 1LUT, 2LUT, 3LUT up to (N-1)-LUT smaller LUTs. In Fig-16A, K is equal to 2. The N-LUT is

constructed as a K-LUT tree, staged with K-LUTs, where  $2^K$  outputs from a first stage feed as LUT values to each of next stage. Each K-LUT has  $2^K$  LUT values and K inputs. There is a  $2^K$  reduction in the number of K-LUTs from one stage to the next. The last K-LUT has only one output. Each K-LUT (such as 1601) in turn is comprised of one or more 1LUTs arranged in one or more stages. The K-LUT is also constructed as a 1LUT tree, staged with 1LUTs, where two outputs of a first stage feed as LUT values to next stage. A secondary K-LUT stage (such as 1605) provides programmability in connecting K-LUTs (from 1601-1604) to form an N-LUT tree. K-LUTs 1601-1604 outputs can bypass K-LUT 1605 to registers. By programming the by-pass option, all K-LUTs can be used independently. A first stage in a secondary K-LUT 1605 comprises 1LUTs having two LUT values that can be configured to be one of two options: programmable logic states (such as 1613 output), or two previous LUT outputs (such as 1611). Except the first stage, every subsequent secondary LUT stages in the N-LUT may have K-LUTs comprising a first stage with this programmable capability. When LUT values are configured as logic states, the N-LUT may compute  $(2^N-1)/(2^K-1)$  independent smaller K-LUT functions. When all secondary LUT values are configured as outputs from previous LUTs, and the K-inputs in each stage is made common to all K-LUTs in that stage, the K-LUT may be used to construct one N-LUT logic function. When all the K-LUT inputs are not made common to all the K-LUTs in that stage, a logic function with more than N-inputs may fit into an N-LUT tree. This hierarchical K-LUTs arrangement is called a LUT macrocell circuit. The LUT macrocell provide programmability to combine multiple smaller LUTs to one larger LUT, or implement logic in smaller LUT form.

The circuit in Fig-16B is only different to that in Fig-16A on the method of choosing inputs to programmable 2LUT 1625. Both A and B inputs have the capability of being selected from external inputs V, X, Y & Z, or prior LUT outputs I<sub>1</sub>, I<sub>2</sub>, I<sub>3</sub> & I<sub>4</sub>. The programmable look up table (LUT) macro-cell circuit for an integrated circuit in Fig-  
5 16B, comprises: a plurality of LUT devices 1621-1625; each said LUT device having an output (such as I<sub>1</sub>-I<sub>4</sub>, F), at least one input (such as A-K), and at least two LUT values; and a programmable means (such as MUX 1651) of selecting inputs to at least one of said LUT devices from one or more other LUT device outputs and external inputs; and a programmable means of selecting LUT values to at least one of said LUT device (such as  
10 1625) from one or more other LUT device outputs and configurable logic states. The crossed-circles show memory bits that need programming to customize the LUT functions. The Silicon consumption for SRAM cells is reduced as demonstrated by the incorporated references.

A programmable macro look up table (macro-LUT) circuit in Fig-16B for an  
15 integrated circuit, comprises: a plurality of LUT circuits (1621-1625), each of said LUT circuits comprising a LUT output, at least one LUT input, and at least two LUT values; and a programmable means (such as 1651) of selecting LUT inputs to at least one of said LUT circuits from one or more other LUT circuit outputs and external inputs, and selecting LUT values to at least one of said LUT circuits (such as 1625) from one or  
20 more other LUT circuit outputs and configurable logic states, said programmable means further comprised of two selectable manufacturing configurations, wherein: in a first selectable configuration, a random access memory circuit (RAM) is formed, said memory circuit further comprising configurable thin-film memory elements; in a second selectable

configuration, a hard-wire read only memory circuit (ROM) is formed in lieu of said RAM, said ROM duplicating one RAM pattern in the first selectable option.

A 5-input LUT (5LUT) can be easily constructed with the method presented in Fig-16. The four circuits 1601-1604 can be replaced by four conventional 3LUTs. The  
5 four outputs can be fed as shown in Fig-16 into the programmable 2LUT. Similarly a 6LUT macrocell can be constructed by constructing four conventional 4LUTs in the first stage in Fig-16. The outputs from 4LUTs are then fed to the programmable 2LUT as shown in Fig-16. Two programmable 3LUT versions are shown in Fig-17A and Fig-17B. In Fig-17A, six 1LUTs as discussed in Fig-13A are combined as shown. In Fig-17B,  
10 seven 1LUTs as discussed in Fig-13A are combined in two stages as shown. A 6LUT macrocell can be constructed by combining six conventional 3LUTs with either of the two programmable 3LUTs shown in Fig-17A and Fig-17B. A programmable look up table (LUT) circuit in Fig-17A for an integrated circuit, comprises: N primary inputs (such as A, B, C), wherein N is an integer value greater than or equal to one, and each  
15 said N inputs received in true and compliment logic levels; and  $2^N$  secondary inputs (such as  $I_1$ - $I_8$ ); and  $2^N$  LUT values, each said LUT values comprising a programmable method to select between one of said secondary inputs (such as  $I_1$ - $I_8$ ) or a configurable logic state (such as one of 1701-1708).

The efficiency of these LUT macrocells in Silicon utilization can be demonstrated  
20 with the 4-variable truth table and the logic function shown in Fig-18A. It realizes a function that lends to truth table logic reduction. A 1LUT gate realization of the function is shown in Fig-18B. It uses only four 1LUTs. The same function is ported to a 4LUT shown in Fig-18C. There are 15 equivalent 1LUTs in the 4LUT, and all are required to

implement the function. The 4LUT is seen to occupy 3.75x more pass-gate Silicon in this example compared to an ideal implementation shown in Fig-18B (without counting the programmable memory bits required to set the LUT values). If we use the 4LUT macrocell shown in Fig-16 which provides 2LUT divisibility, this function can be implemented as shown in Fig-18D. The bit polarity required to achieve the desired functionality are shown next to each bit in Fig-18D. That allows two 2LUTs 1803 and 1804 to be used for other 2-input logic functions. Those outputs can be taken out to registers via the by-pass circuitry. The macrocell shown in Fig-16 can be partitioned into 2LUTs by design and used as five 2LUT blocks. It uses an equivalent of 21 1LUT gates, compared to 15 for the 4LUT in Fig-18C. Column-4 in Fig-4 shows that 4LUT on the average is only 36% efficient compared to 2LUTs at fitting logic. Accounting for 21/15 inefficiency for the larger Si foot-print in the 4LUT macrocell in Fig-16, it is still ~2X more efficient at fitting an average logic design in 2LUT pieces.

Each of the circuits described in Fig-11 thru Fig-17 provides a programmable means to configure the LUT macrocell. Said programmable means comprises a memory circuit fabricated with two selectable manufacturing configurations. In a first selectable configuration a RAM circuit is formed to provide said LUT user re-programmability. In a second selectable configuration a ROM circuit is formed in lieu of one specific RAM pattern to provide identical LUT programmability.

New programmable LUT circuits are described for use in large and fine geometry FPGA devices. As the logic density increases, there is a need to add more LUTs into a logic block, and increase the LUT size. Both inhibit the efficiency of Silicon utilization when porting logic synthesized to an ASIC flow. Compared to 2LUT based logic blocks,

4LUTs are seen to be only 36% efficient, while 7LUTs are only 7% efficient. The new LUT circuits disclosed herein make use of additional programmable elements inside the large LUT structure, enabling sub-division of LUTs. A complex design can be fitted as a single larger logic LUT or as many smaller logic LUT pieces: both maximizing the Silicon utilization. A 2LUT divisible 4LUT macro-cell shown in Fig-16A provides a 2X improvement in logic packing compared to hard-wired 4LUT logic elements. The increased memory content is justified by a 3-dimensional thin-film transistor module integration that allows all configuration circuits to be built vertically above logic circuits. These memory circuits contain memory elements that control pass-gates constructed in substrate Silicon. The TFT layers are fabricated above a contact layer in a removable module, facilitating a novel method to remove completely from the process. Configuration circuits are mapped to a hard-wire metal links to provide the identical functionality in the latter. Once the programming pattern is finalized with the thin-film module, and the device is tested and verified for performance, the TFT cells can be eliminated by hard-wire connections. Such conversions allow the user a lower cost and more reliable end product. These products offer an enormous advantage in lowering NRE costs and improving TTS in the ASIC design methodology in the industry.

Although an illustrative embodiment of the present invention, and various modifications thereof, have been described in detail herein with reference to the accompanying drawings, it is to be understood that the invention is not limited to this precise embodiment and the described modifications, and that various changes and further modifications may be effected therein by one skilled in the art without departing from the scope or spirit of the invention as defined in the appended claims.